



# Immuno-informatics based approaches to design a novel multi epitope-based vaccine for immune response reinforcement against Leptospirosis

Majid Valadi<sup>a,1</sup>, Ahmad Karkhah<sup>b,c,1</sup>, Vijay Kumar Prajapati<sup>d</sup>, Hamid Reza Nouri<sup>c,\*</sup>

<sup>a</sup> Clinical Biochemistry Research Center, Basic Health Sciences Institute, Shahrekord University of Medical Sciences, Shahrekord, Iran

<sup>b</sup> Student Research Committee, School of Medicine, Babol University of Medical Sciences, Babol, Iran

<sup>c</sup> Cellular and Molecular Biology Research Center, Health Research Institute, Babol University of Medical Sciences, Babol, Iran

<sup>d</sup> Department of Biochemistry, School of Life Sciences, Central University of Rajasthan, Bandarsindri, Kishangarh, Ajmer, Rajasthan, India

## ARTICLE INFO

### Keywords:

Leptospirosis  
Immunoinformatics  
Multi-epitope vaccine  
Reverse vaccinology

## ABSTRACT

Leptospirosis is known as a zoonotic disease of global importance originated from infection with the spirochete bacterium *Leptospira*. Although several leptospirosis vaccines have been tested, the vaccination is relatively unsuccessful in clinical application despite decades of research. Therefore, this study was conducted to construct a novel multi-epitope based vaccine against leptospirosis by using Hap1, LigA, LAg42, SphH and HSP58 antigens. T cell and IFN gamma epitopes were predicted from these antigens. In addition, to induce strong CD4<sup>+</sup> helper T lymphocytes (HTLs) responses, Pan HLA DR-binding epitope (PADRE) and helper epitopes selected from Tetanus toxin fragment C (TTFC) were applied. Moreover, for boosting immune response, Heparin-Binding Hemagglutinin (HBHA), a novel TLR4 agonist was added to the construct as an adjuvant. Finally, selected epitopes were linked together using EAAAK, GPFG, AAY and HEYGAEALERAG linkers. Based on the predicted epitopes, a multi-epitope vaccine was construct with 490 amino acids. Physico-chemical properties, secondary and tertiary structures, stability, intrinsic protein disorder, solubility, and allergenicity of this vaccine construct were assessed by applying immunoinformatics analyses. A soluble, and non-allergic protein with a molecular weight of 53.476 kDa was obtained. Further analyses validated the stability of the chimeric protein and the predicted epitopes in the chimeric vaccine indicated strong potential to induce B-cell and T-cell mediated immune response. Therefore, immunoinformatics analysis showed that the modeled multi-epitope vaccine can properly stimulate the both T and B cells immune responses and could potentially be used for prophylactic or therapeutic usages.

## 1. Introduction

Leptospirosis is known as a zoonotic disease of global importance originated from infection with the spirochete bacterium *Leptospira* (Vinetz, 2001). *Leptospira* are highly motile, obligate aerobic spirochetes that mimic features of both gram-positive and gram-negative bacteria (Palaniappan et al., 2007). *Leptospira* are about  $0.25 \times 6 - 25 \mu\text{m}$  in size and can easily pass through  $0.45 \mu\text{m}$  filters. The genetically based classification of *Leptospira* indicates that there are at least 19 species that seven of these species were introduced as main causes of leptospirosis including *L. interrogans*, *L. borgpetersenii*, *L. santarosai*, *L. noguchii*, *L. weilii*, *L. kirschneri* and *L. alexanderi* (Evangelista and Coburn, 2010). Leptospirosis was firstly presented as Weil's syndrome (Adler et al., 2010). Leptospirosis has a worldwide distribution and has also emerged as a disease of the adventure

traveller, especially affecting humans who are frequently in contact with rodents, pets or polluted water (Trueba et al., 2004; Waitkins, 1986). This spiral-shaped bacterium can the ability of penetrating the mucosa or an open skin. After penetrating, the bacterium can lead to serious clinical manifestation (Schmid et al., 1986). These clinical presentations comprise hepatitis, nephritis, meningitis, pneumonitis, and pancreatitis and ultimately death (Christova et al., 2003; Katz et al., 2001). The interaction of leptospires with pathogen recognition receptors is a major issue in immune responses against *Leptospira* and in immunopathology (Fraga et al., 2011). Pathogenic leptospires are able to evade the host immune system, circulate in the blood and spread into tissues. The rapid dissemination of pathogenic *leptospira* to kidney fibroblasts lead to induction of apoptosis or programmed cell death and result in renal failure (Merien et al., 1997). Therefore, inadequate function of immune response against pathogenic *leptospira* is a new

\* Corresponding author.

E-mail address: [nourih851@gmail.com](mailto:nourih851@gmail.com) (H.R. Nouri).

<sup>1</sup> Contributed equally to this work.

challenge to develop an effective approach to reinforce immune system responses against leptospirosis. Recently, molecular and cellular studies on leptospirosis have been focused on the field of leptospirosis vaccines (Adler, 2015). Although vaccines to prevent human leptospirosis are available in some countries, several problems confront the development of a vaccine to prevent human leptospirosis including an unacceptable side effect profile of killed bacterial vaccines, short-term and incomplete protection of killed bacteria vaccines, the locally varying patterns of *Leptospira*, theoretical potential for inducing autoimmune disease such as uveitis and, finally an incomplete knowledge of mechanisms of protective immunity against leptospiral infection (Koizumi and Watanabe, 2005).

To design a *Leptospira* vaccine, conserved protein components of different serovars of *Leptospira* are established as ideal targets. Among vaccine-related genes were founded through the web based VIOLIN vaccine database and previously experimental studies, hemolysis-associated protein-1 (Hap1), leptospiral immunoglobulin (Ig)-like proteins A (LigA), LAg42, hemolysin SphH and heat-shock protein 58 (HSP58) were considered as priority antigens for developing a subunit vaccine. The Hap1 (also known as LipL32) was determined as a protein expressed in pathogenic *Leptospira*. The Hap1 was used as an immunogen through Adenovirus delivery and induced significant protection against a virulent heterologous *Leptospira* challenge in vaccinated animal models (Branger et al., 2001). Also DNA vaccines expressing Hap1 have been shown cross-protective effect with pathogenic strains of *Leptospira* (Branger et al., 2005). Hence, it could be a good candidate for development of new vaccines. In addition, Lig proteins are a fraction of leptospiral surface components that binds to extracellular matrix (ECM). Therefore, these proteins may be involved in bacterial attachment to host tissues and facilitate the bacterial colonization process. LigA contains a lipoprotein signal peptide that followed by 12–13 series of Ig-like domains. The first six domains of LigA along with a portion of the seventh domain are 99% identical with LigB (Lin et al., 2009; Lin et al., 2010). Currently, several studies have shown that LigA subunit or DNA vaccines generated from different serovars could induce protective immune responses in rodent models and protected them from homologous challenge infection (Koizumi and Watanabe, 2004; Palaniappan et al., 2006; Silva et al., 2007). The novel antigen LAg42 was first introduced by Koizumi and et al. The LAg42 gene was conserved among various pathogenic strains of *Leptospira* but not among non-pathogenic leptospires. The LAg42 is an inner membrane protein with molecular weight 42 kDa that its C-terminal is immunogenic. Also *in vivo* experiments suggest that LAg42 may be a good target for protective immune response (Koizumi and Watanabe, 2003). Furthermore, it has been shown that hemolysins are virulence factors for the pathogenesis of leptospirosis. The SphH is a novel *Leptospira* hemolysin determined among the pathogenic *Leptospira* and act as a pore-forming protein. The SphH is highly conserved among pathogenic leptospires and directly caused membrane damage of cells that indicated its potential role in pathogenesis of leptospirosis (Lee et al., 2002). Additionally heat shock proteins (HSPs) as molecular chaperons are highly conserved molecules that protect cells from various stimuli and act as major antigens of various microbial pathogens. They significantly induce humoral and cellular immune responses. In this context, HSP58 antigen is recognized by antibodies in sera of patients infected with several serovars of leptospires. More *in vitro* experiments introduced a specific region of amino acids in the HSP58 for its immunoreactivity. Hence, it can apply as an important target for humoral immune responses induction during leptospirosis (Flannery et al., 2001; S. H. Park et al., 1999). Therefore, to develop efficient subunit vaccine against leptospirosis, these immunogenic proteins were targeted.

Considering recent advances in the field of vaccine development using bioinformatics and immunoinformatics approaches (Karkhah et al., 2017; Saadi et al., 2017), these problems seems resolvable in the development of a vaccine to prevent human leptospirosis. Considering the benefits of multi-epitope vaccines including high immunogenicity,

stability in various conditions, and high specificity as well as affordable production, in this study we developed a multi-epitope protein from immune-dominant peptides of Hap1, LigA, LAg42, SphH and HSP58 antigen with appropriate linkers using bioinformatics tools. Modeling of protein structure and codon optimization for expression in *Escherichia coli* was carried out and the potential of the developed vaccine for inducing the immune responses were evaluated.

## 2. Materials and methods

### 2.1. Sequence retrieval and homology analyses

In the first step, Uniprot Knowledgebase data was used to retrieve the amino acid sequences of Hap1 (Accession no. Q79B72), LigA (Accession no. C7SAC0), LAg42 (Accession no. AB091092), SphH (Accession no. O34095), HSP58 (Accession no. P61439) antigens, tetanus toxin fragment C (Accession no. 1A8D\_A), and HBHA of *Mycobacterium tuberculosis* (Accession no. ZP\_07011362.1) in FASTA format. L. interrogans was targeted for antigen selection. The selected antigens may have variation in amino acids sequences between seven species of *Leptospira* were responsible for leptospirosis, therefore multiple sequence alignments (MSA) were performed on deposited sequence of selected antigens to assess the conserved regions. The MSA was performed by CLUSTALW at <https://www.genome.jp/tools-bin/clustalw> or National Centre for Biotechnology Information (NCBI) at <http://www.ncbi.nlm.nih.gov/Blast>. However, all subsequent analyzes were performed on conserved sequences.

### 2.2. T-cell epitopes Prediction

Potential T cell epitopes in Hap1, LigA, LAg42, SphH and HSP58 proteins were predicted by using CTLpred online server (Bhasin and Raghava, 2004). This server is mainly based on elegant machine learning techniques such as Vector Machine (SVM) and Artificial Neural Network. CTL epitopes were also predicted in consensus and combined approaches for more specific and sensitive results.

### 2.3. Interferon-gamma inducing epitopes prediction

To obtain a vaccine with strong ability of the immune system reinforcement, interferon-gamma inducing epitopes were also studied by using the IFNepitope server at <http://crdd.osdd.net/raghava/ifnepitope/> (Dhanda et al., 2013). This web server works based on a dataset including IFN-gamma inducing and non-inducing MHC class II binders through different algorithms, such as machine learning technique, motifs-based search, and hybrid approach.

### 2.4. MHC binding epitopes predictions

ProPred online server at <http://www.imtech.res.in/raghava/propred/> was used to predict MHC class I- binding peptides. This online server provides quantitative matrices for 47 MHC Class-I alleles. The maximum accuracy of ProPred server is 75% at default threshold. Another server IEDB MHC I prediction tool at <http://tools.iedb.org/mhci/nthat> was used for the prediction of MHC class I- binding peptides. MHC-II peptide binders were predicted by RANKPEP from protein sequences at <http://bio.dfci.harvard.edu/RANKPEP>. This server utilizes position-specific scoring matrix (PSSM) to predicts peptides.

### 2.5. Determination of conserved domain

NCBI's Conserved Domain Database (CDD) was applied to predict conserved structural and functional amino acid sequences at <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>. This data bank can identify conserved domains in protein structures using RPS-BLAST. Moreover, predicted epitopes were subjected for analysis of epitope

conservancy at <http://tools.iedb.org/conservancy/> (Bui et al., 2007).

## 2.6. Vaccine construct design and chimeric gene optimization

In this stage, Hap1, LigA, LAg42, sphH and HSP58 proteins were used for selection of potential T cell and interferon-gamma inducing epitopes. In addition, TTFrC and PADRE were applied as helper epitopes. In addition, Heparin-Binding Hemagglutinin (HBHA) a novel TLR4 agonist was added to the construct as an adjuvant for boosting immune response. In the next stage, each peptide segment that was derived from the helper antigens, the main antigens and HBHA protein was linked to each other by proper amino acid linkers. Different features of the aforementioned multi-epitope protein were calculated by different online servers. In addition, JCAT server (<http://www.jcat.de>) and DNA2 software were used to obtain the DNA sequence of designed protein by reverse translation and its adaptation to *E. coli* codon usage (Grote et al., 2005).

## 2.7. mRNA secondary structure prediction

Mfold software was applied for prediction of RNA secondary structure. Mfold software can provide the prediction of true positive base pairs and the structures with minimum  $\Delta G$  thermodynamically (Zuker, 2003).

## 2.8. Physico-chemical parameters analysis

ProtParam online server (<http://us.expasy.org/tools/protparam>) was used to evaluate physico-chemical parameters including the amino acid composition, theoretical isoelectric point, molecular weight (MW), in-vitro and in-vivo half-life, aliphatic index, instability index, extinction coefficient, grand average of hydropathicity (GRAVY) and the total number of positive and negative residues (Wilkins et al., 1999).

## 2.9. Prediction of secondary and tertiary structure

GOR IV and PSIPRED servers were applied for prediction of the secondary protein structure at <http://expasy.org/tools/gor4.html> and <http://bioinf.cs.ucl.ac.uk/psipred/>, respectively (McGuffin et al., 2000; Sen et al., 2005). The 3D structure was predicted by using the I TASSER and Phyre2 (Kelley and Sternberg, 2009; Zhang, 2008). The stereochemical validity of the protein structure and possible errors in 3D models were determined in PROCHECK program by Ramachandran plot at [http://www.ebi.ac.uk/thornton\\_srv/software/PROCHECK](http://www.ebi.ac.uk/thornton_srv/software/PROCHECK) (Lovell et al., 2003; Wiederstein and Sippl, 2007).

## 2.10. Refinement of the 3D modeled structure

The refinement of modeled structure was performed by applying GalaxyRefine server at <http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>. Briefly, modeled structure by using Phyre2 was subjected to refine whole protein. GalaxyRefine performs repeated structure perturbation and subsequent overall structural relaxation by molecular dynamics simulation (H. Park and Seok, 2012).

## 2.11. Prediction of intrinsic protein disorder

DisEMBL 1.5 at <http://dis.embl.de/> and IUPred at <http://iupred.enzim.hu/pred.php> were applied to assess intrinsic protein disorder and unstructured regions within a protein. DisEMBL is a public web server for predicting disorder in proteins. In addition, DISOPRED3 was used to predict intrinsic protein disorder and unstructured regions in our constructed vaccine (Jones and Cozzetto, 2015).

## 2.12. Prediction of B-cell epitopes

Two online servers were applied to predict both linear and conformational B cells epitopes. Linear B-cell epitopes were predicted by BCPREDS server at <http://ailab.cs.iastate.edu/bcpreds/predict.html> using subsequence kernel, as a novel method, with specificity of 74.5% and conformational B cell epitopes were determined from 3D protein structure using Discotope server at <http://www.cbs.dtu.dk/services/DiscoTope/> (Kringelum et al., 2012; Saha and Raghava, 2006b).

## 2.13. Evaluation of allergenicity, antigenicity, sequence similarity and epitope conservancy analysis

In order to determine protein allergenicity with a high accuracy, based on six different approaches, AlgPred web server at <http://www.imtech.res.in/raghava/algpred/> was applied (Saha and Raghava, 2006a). The accuracy of AlgPred server is about 85% at a threshold -0.4. Antigenicity was also analyzed applying VaxiJen v2.0 server at <http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html> (Doytchinova and Flower, 2007). Also, for evaluation the proposed multi-epitope vaccine similarity with the human proteins, firstly amino acid sequence of the proposed vaccine was submitted to the Basic Local Alignment Search Tool (BLAST) in NCBI. After that, the similarity between the suggested sequences and multi-epitope vaccine was determined by the multiple sequence alignment (MSA) through Clustal Omega in EMBL-EBI database. In addition, epitope conservancy analysis was performed at <http://tools.iedb.org/conservancy/> to computes the degree of conservancy of the selected epitopes in proposed vaccine within protein sequence of the other pathogenic species of *Leptospira*.

## 2.14. Prediction of protein solubility

Propensity of protein solubility was predicted by SOLpro server at <http://scratch.proteomics.ics.uci.edu/>. SOLpro predicts the propensity of a protein to be soluble upon overexpression in *E. coli* using a two-stage SVM architecture based on multiple representations of the primary sequence (Magnan et al., 2009).

## 2.15. Ligand binding site exploration and molecular docking

For protein-protein docking of the HBHA part of the vaccine as a ligand and the TLR4 as the receptor, ClusPro at <http://nrc.bu.edu/cluster> was used. ClusPro represents the first fully automated, web-based program for the computational docking of protein structures.

# 3. Results

## 3.1. Primary analysis of the retrieved sequences

In the present study, Hap1, LigA, LAg42, sphH and HSP58 protein sequences were applied to immunoinformatic analysis to determine CD8<sup>+</sup> CTL, IFN gamma and MHC binder's epitopes. Furthermore, TTFrC and PADRE were used for selection of CD4<sup>+</sup> helper epitopes. Since the vaccine needs the strong induction of T cell immune responses, highly conserved regions of HBHA were used as adjuvant. The signal sequences were omitted from all constructed structures.

## 3.2. Prediction of IFN- $\gamma$ and CTL epitopes

The high score and suitable CTL epitopes were predicted using CTLpred server from Hap1, LigA, LAg42, sphH and HSP58 protein sequences (Tables 1). Five regions were selected from Hap1, LigA, LAg42, SphH and HSP58 protein sequences which were subjected to in-silico analysis in order to collect the best immunoreactive epitopes. In addition, interferon-gamma inducing epitopes were also predicted by using the IFNepitope server which can predict IFN-gamma inducing and non-

**Table 1**  
Final IFN- $\gamma$  and CTL epitopes selected from Hap1, LigA, LAg42, sphH and HSP58 antigens.

Selected Antigen	Start Position	End Position	Sequence	IFN epitope	Identity	Antigenicity score
Hap1	63	78	KPGQAPDGLVDGNKK	+	100%	0.92
	72	87	VDGNKKAYLYVWIP	+	100%	1.8
LigA	11	26	DNSNSDITDQVTWDS	+	100%	0.49
	18	33	TDQVTWDSNTDILS	+	100%	0.14
LAg42	54	69	QKLVVVGCAERYPD	+	100%	0.19
	59	74	VGCAERYPDNIHSE	+	100%	0.61
sphH	19	34	KSLIGFYCLFMFFLN	+	100%	1.80
	28	43	FMFFLNCLPDKQKEH	+	100%	0.90
HSP58	42	57	FGAPTITKDGVTVAK	+	100%	0.74
	69	84	AQMVKEVSTKTNDVA	+	100%	0.96

inducing MHC class II binders.

3.3. Defining MHC-I and MHC-II binding epitopes

High ranked peptide regions for presenting by MHC class I over several alleles (MHC-2Kb, MHC-2Dd, MHC-2Kd, MHC-2Kk and MHC-2Ld) were predicted by applying ProPred server. To confirm the defined peptides, the IEBD server was also used. Moreover, the high score binding regions to HLA class II were selected by RANKPEP online server.

3.4. Design and construction of multi-epitope peptide vaccine

The final protein sequence was composed of 490 amino acids as shown in Fig. 1. As the first domain, HBHA was selected in the N terminal of the construct as an adjuvant. As the second domain, four regions from TTFrC (Table 2) and PADRE were applied as CD4<sup>+</sup> helper epitopes. The third domain was constructed by CTL and IFN-gamma inducing epitopes selected from Hap1, LigA, LAg42, sphH and HSP58 protein sequences. Finally, proper linkers were used to make a linkage between favorite peptide sequences. As mentioned above, to obtain codon-optimized DNA sequence of the constructed protein, JCAT online server was employed. JCAT contributes to optimized cloning and expression in *E.coli* as a host. High-level and optimum expression of chimeric sequence are more than 0.8 according to Codon adaptation index (Schnoor et al., 2015). CAI in the adapted sequence was 0.97. GC-Content of the improved sequence was also 51.63 % and GC-Content of *E. coli* was 50.7. Graphical view of codon usage in optimized chimeric gene was shown in Fig. 2.

3.5. mRNA structure prediction

Free energy related to whole mRNA structure and its 5' end was obtained using Mfold online service and its details are shown in Table 3. The minimum free energy of secondary RNA structure was  $\Delta G = -529.00$  kcal/mol without hairpin nor pseudoknot at 5' side (Figs. 3 and 4).

3.6. Physico-chemical parameter, allergenicity, antigenicity and solubility

The molecular weight and theoretical isoelectric point value of the chimeric protein were 53.476 kDa and 5.13, respectively. Extinction

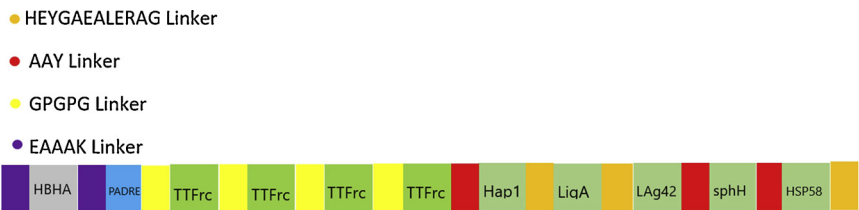
**Table 2**  
TTFrC MHC-II binding peptides determined by RANKPEP.

Epitope sequence	Start position	End position
NDIISDISGFNSSVITYPDACLVPGINGKAIHLVNNE	40	66
IEYNDMFNNFTVSFWLRVPKVSASHLEQYGT	78	108
YVSIDKFRIFCKALNPKEIEKLYTSYLS	220	247
LRVGYNAPGIPLYKKMEAVKLRDLK	362	386

coefficients in units of  $M^{-1} cm^{-1}$  at 280 nm measured in water were 55030 and with absorption 0.1% ( = 1 g/l) 1.238, assuming all pairs of Cys residues form cystines. The estimated half-life was reported 30 hours for mammalian reticulocytes in vitro, > 20 hours for yeast in vivo, and > 10 hours for *E. coli* in vivo. The instability index (II) was computed to be 27.37, thereby classifying the constructed protein as stable. Aliphatic index and Grand average of hydropathicity (GRAVY) were obtained to be 82.76 and -0.353, respectively. Obtained result from ALgPred using the hybrid method indicated that the protein sequence did not contain experimentally proven IgE epitope and the vaccine was non-allergen. In addition, antigenicity was also analyzed and the overall prediction for antigenicity was reported 0.67 that suggested the proposed multi-epitope molecule can be an appropriate antigen. In contrast, human protein similarity was also assessed by performing protein BLAST and its observation confirms that there was the lowest possible similarity between reported human proteins and proposed vaccine molecule (Fig. 5). This observation assures that the designed molecule will be recognized as foreign and immunogenic molecule, which will be capable to induce the humoral/cell mediated immune response. In addition, the obtained results of epitope conservancy analysis showed the high similarity in other strains, and then we can conclude that the epitopes which we have used for vaccine designing will be capable to induce same immune response against all pathogenic strains irrespective of their geographical distribution. Furthermore, the constructed vaccine was soluble upon overexpression in *E. coli* with probability of 0.99.

3.7. Secondary and tertiary structure prediction

The secondary structure of the protein that was predicted by PSIPRED and GOR4 secondary structure prediction servers is composed of 52.04 % alpha helix, 13.88% extended strand and 34.08% random



**Fig. 1.** Schematic image of vaccine construct consists of selected peptides from Hap1, LigA, LAg42, sphH and HSP58 antigens, TTFrC and PADRE as helper epitopes and HBHA as an adjuvant fused together by appropriate motifs for expression in *Escherichia coli*.



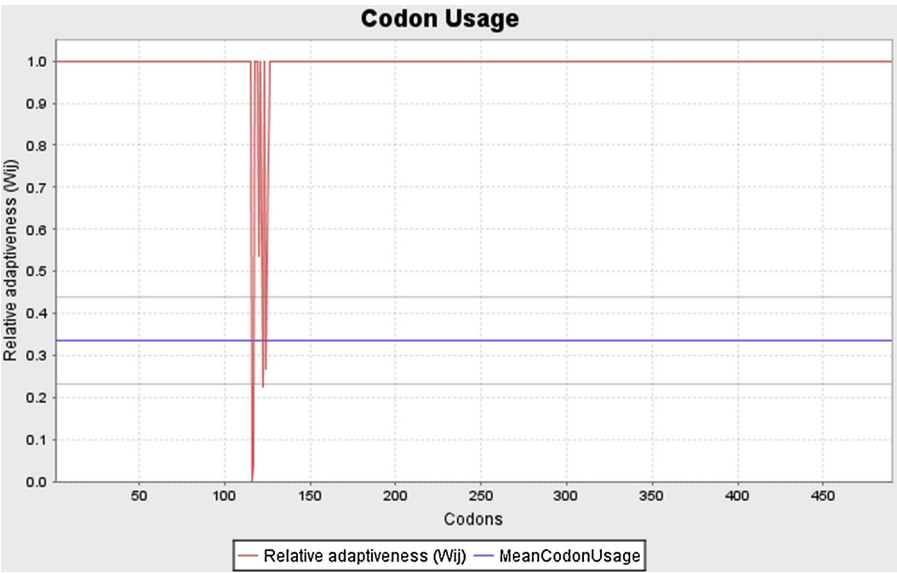


Fig. 2. Graphical view of codon usage in optimized chimeric gene.

**Table 3**  
Free energy details related to 5' end of chimeric gene mRNA structure by mfold web server.

Structural element	Free energy (kcal/mol)	Base pair
External loop	-1.80	17 ss bases & 2 closing helices.
Stack	-0.90	External closing pair is U <sup>603</sup> -A <sup>1461</sup>
Stack	-2.40	External closing pair is U <sup>604</sup> -A <sup>1460</sup>
Stack	-3.30	External closing pair is C <sup>605</sup> -G <sup>1459</sup>
Helix	-6.60	4 base pairs.

coil as presented in Figs. 6 and 7. Using I-TASSER and Phyre2 services, it was shown that there are three main domains in tertiary structure. Phyre2 server indicated that 33% of residues were modelled at > 90% accuracy.

3.8. Refinement and validation of the 3D refined structures

The obtained 3D structures were submitted for the refinement process by GalaxyRefine program. The GalaxyRefine server also induced five 3D refined models, which all entered the next step. The qualities of the 3D structures were improved after refinement according

to the results of the ProSA-web and Ramachandran Plot (Fig. 8). As shown in Fig. 9, validation of protein 3D model by Ramachandran plot before and after refinement indicated that in initial model, 356 (73.0%), 62 (12.7%) and 70 (14.3%) of residues were located in favored, allowed and outlier regions, respectively. In the refined model, 435 (89.1%), 39 (8.0%) and 14 (2.9%) of residues were located in favored, allowed and outlier regions, respectively.

3.9. Intrinsic disorder prediction

Disordered regions were identified and annotated by DisEMBL online service. Amino acids in the input sequence are considered disordered by Loops/coils definition in the following regions: 41-71, 361-381. Disordered regions were also shown in Fig. 10.

3.10. Prediction of B-cell epitopes

As mentioned in methods, the final designed protein was applied to the prediction of continuous and discontinuous B-cell epitopes by BCPRED and Discotope 2.0 servers, respectively. Predicted linear B-cell epitopes from full length constructed vaccine were shown in Table 4. In addition, epitopes predicted in modeled vaccine based on different parameters of epitopes were shown in Table 5. Moreover, using

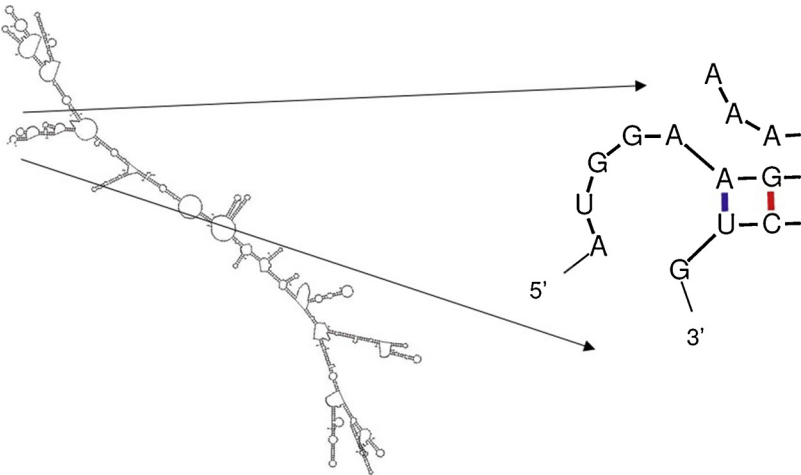
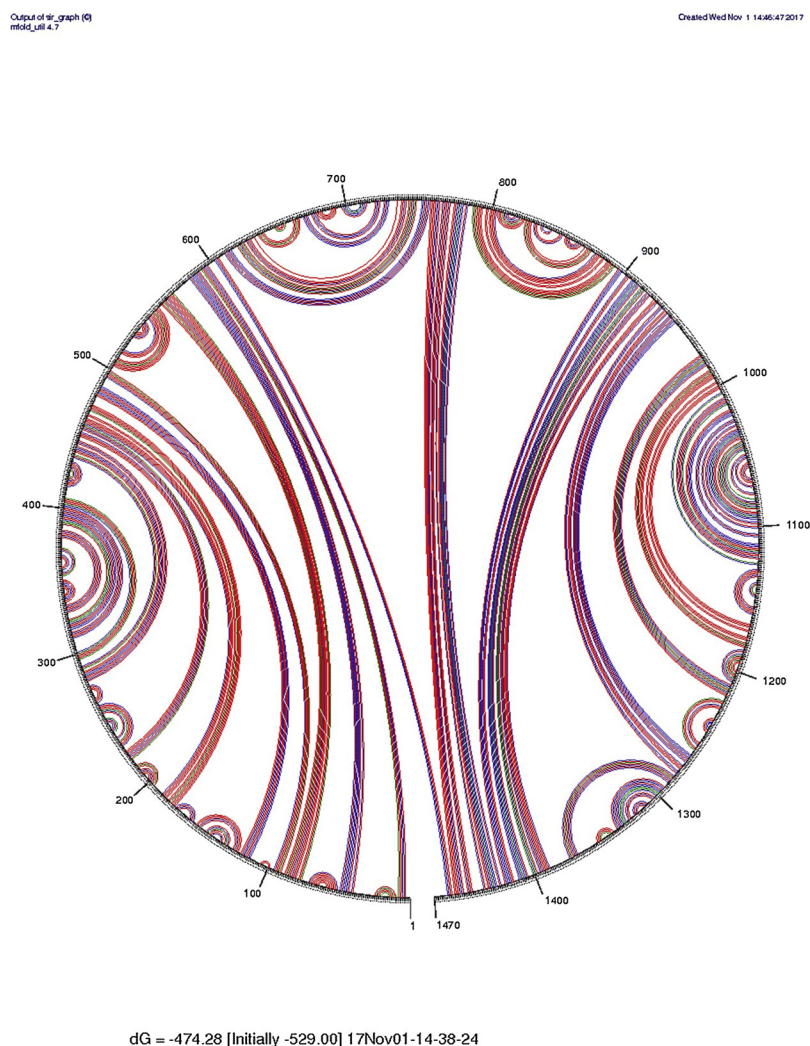


Fig. 3. Prediction of RNA secondary structure of chimeric gene by mfold server. Predicted structure has no hairpin and pseudo knot at 5'site of mRNA.



**Fig. 4.** Circle graph structure. A circle graph is one way to display base pairs of a structure. It also represents a first step to producing a formal structure in which overlap of bases is usually small.

DiscoTope 2.0 server, many conformational B-cell epitopes were identified in the final 3D model.

### 3.11. Exploration of interaction between protein vaccine and TLR4

To evaluate the interaction of the designed vaccine and TLR4, molecular docking of the 3D models of these two proteins was done with ClusPro 2.0. The docked models that represented interactions in mis-allocated segments of either vaccine or TLR4 were excluded, and the best docking model (shown in Fig. 11) was selected based on the residues involved in the interactions.

## 4. Discussion

Leptospirosis is a serious infection disease caused by pathogenic strains of the *Leptospira* spirochetes, which affects not only humans but also animals (Adler et al., 2010). Although some leptospirosis vaccines have been obtained, the vaccination is relatively unsuccessful in clinical application despite decades of research (Wang et al., 2007). There are two basic types of leptospirosis vaccines available, attenuated and inactivated leptospirosis vaccines. However, these two types of vaccines reveal significant safety problems. Recombinant protein vaccines have a great potential against leptospirosis (Koizumi and Watanabe, 2005; Vinetz, 2001). Before production of recombinant protein vaccines against leptospirosis for clinical application, extensive testing is

required. Recombinant protein vaccines must be stable and safe, and have high immunogenicity, stability in various conditions, and high specificity as well as affordable production. Considering recent advances in the field of vaccine development using bioinformatics and immunoinformatics approaches, these issues seems resolvable in the development of a vaccine against human leptospirosis (Karkhah and Amani, 2016; Tourani et al., 2017). Therefore, in this study we developed a novel multi-epitope protein from immune-dominant peptides of Hap1, LigA, LAg42, sphH and HSP58 antigens. For designing our multivalent vaccine using immuno-informatics approaches, five regions from these antigens were selected as T cell and IFN-gamma inducing epitopes. Several lines of evidence introduced IFN-gamma as the signature cytokine of both the innate and adaptive immune system eliciting antiviral, immune regulatory as well as anti-tumor activities. The release of IFN-gamma is the major arm of the Th1 response critical for the control of pathogens such as leptospirosis (Schoenborn and Wilson, 2007). Moreover, to induce strong CD4<sup>+</sup> helper T lymphocytes (HTLs) responses, four regions of TTFrC along with Pan HLA DR-binding epitope (PADRE) were defined as helper epitopes. A key feature of this universally immunogenic epitopes is its capacity to bind to several isotypic and allotypic forms of human MHC class II molecules inducing an increased and long term immune response by increasing the helper T-cell response (Panina-Bordignon et al., 1989). It has been reported that leptospirosis vaccines with adjuvant are more immunogenic than those without adjuvant (Fraser et al., 2007). Therefore, Heparin-



Fig. 5. Multiple sequence alignment between multi-epitope proposed vaccine and human proteins showed the lowest possible similarity.

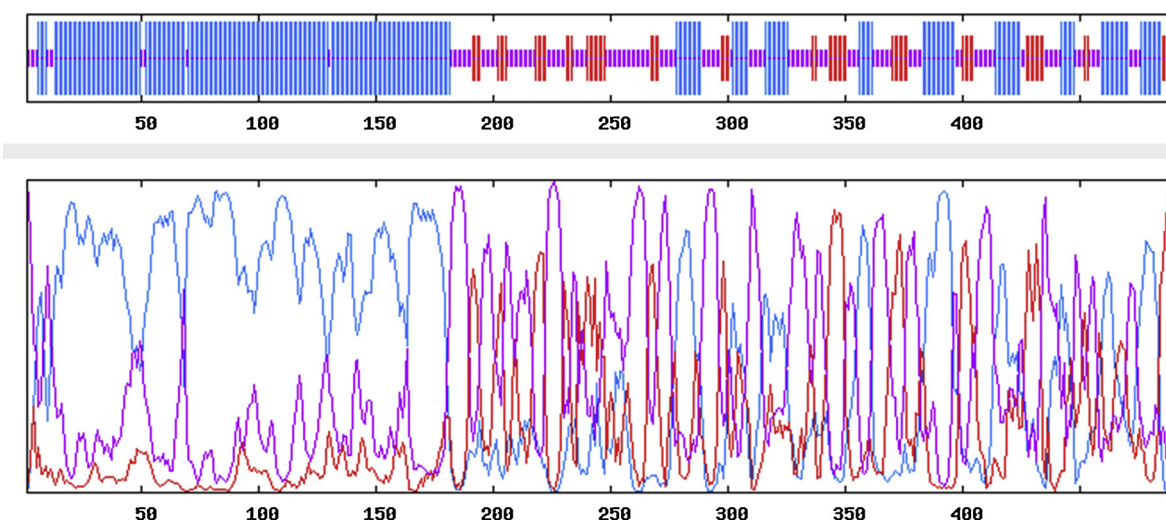


Fig. 6. GOR IV server results indicated that constructed vaccine includes 52.04% alpha helix (H), 13.88% extended strand and 34.08% random coil (C) in secondary structure.

Binding Hemagglutinin (HBHA), a novel TLR4 agonist was added to the construct as an adjuvant. The final vaccine consisted of three domains: HBHA as an adjuvant, helper epitopes (TTFrC and PADRE) and antigenic epitopes. GPGPG, AAY and HEYGAELERAG cleavable linkers were used to separate these three domains from each other in order to enhance epitope presentation. These linkers have dual roles in epitope vaccine structure; first, preventing the generation of junctional epitopes (neoepitopes) that is the important concern in designing epitope vaccines, and second, facilitating the immune processing and presentation of HLA-II binding epitopes (Livingston et al., 2002; Meza et al., 2017; Nezafat et al., 2016). In addition, because of functional features of HBHA, as TLR4 agonist, the EAAAK linker was used to link HBHA to N terminal of vaccine construct as an adjuvant to reduce the interaction with other vaccine sections and cause more effective separation (Arai

et al., 2001). In addition, it should be noted that there are some investigations demonstrated that GPGPG spacer construct can be used to induce HTL responses by either polypeptide or DNA immunization to control infection (Livingston et al., 2002). Because the induction of Th lymphocytes are a crucial component of both the humoral and cellular immune response and also the secretion of cytokines by HTL is important in the differentiation of CTL (Chaplin, 2010); hence, these immune-dominant epitopes and linkers were used in our multi- epitope construct.

Therefore, a multi-epitope vaccine was constructed with 490 residues in length. Restriction sites, instability elements, and all the cis-acting sites were removed from the construct which significantly interfering with cloning. In addition, codon optimization was performed to improve transcriptional and translational efficiency and to facilitate

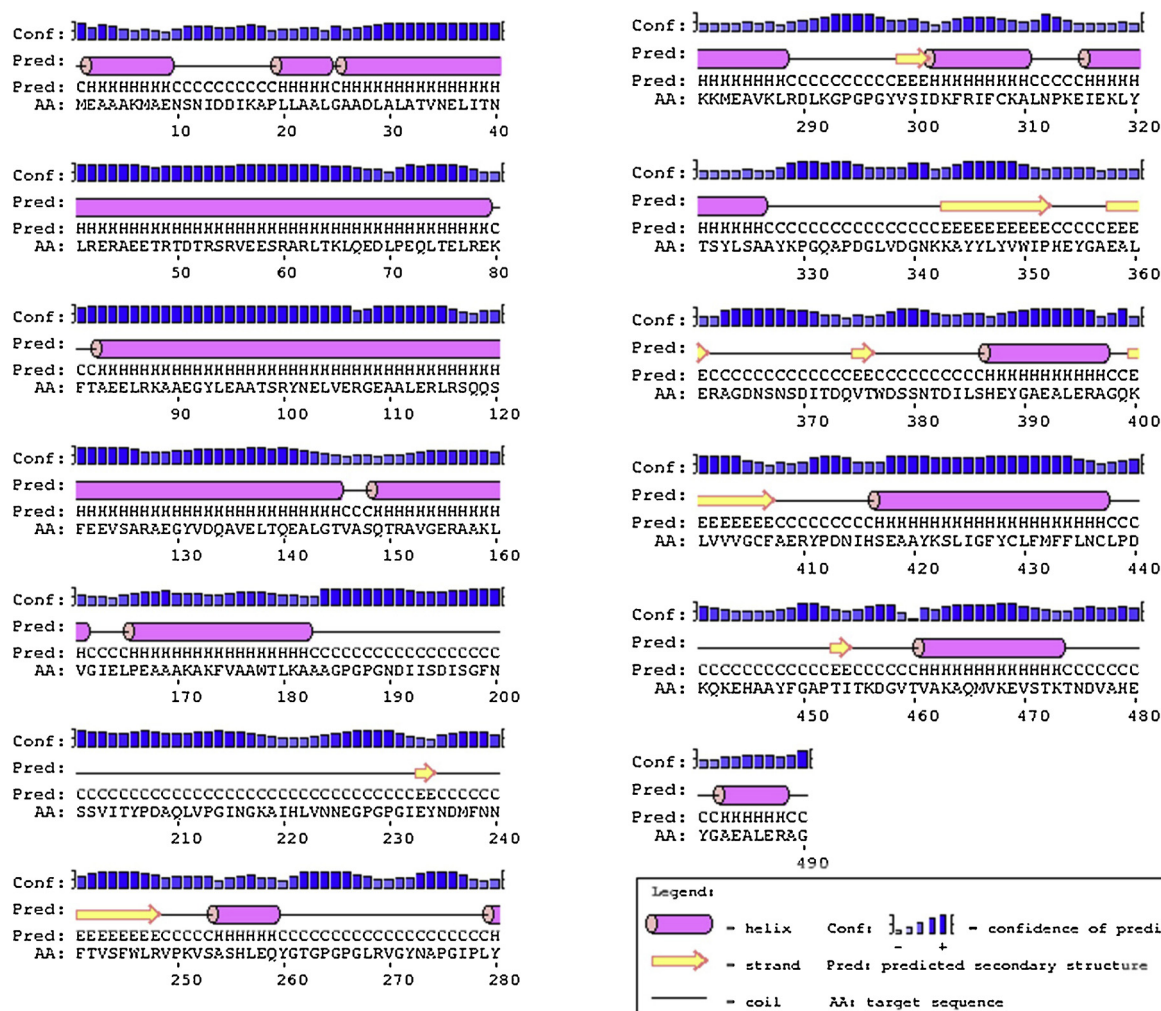


Fig. 7. Graphical result from secondary structure prediction of vaccine using PSIPRED.

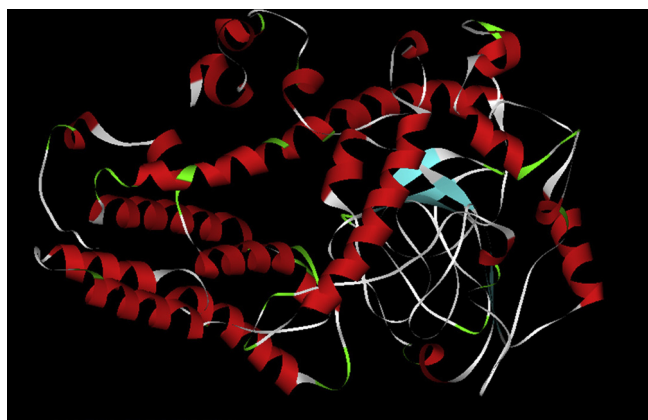


Fig. 8. Refined modeled structure of constructed vaccine using Phyre2 software.

high-level expression of recombinant protein vaccine in *E. coli*. mRNA secondary structure is a significant factor in expression of proteins. The results from mRNA prediction by Mfold server indicated that the mRNA had enough stability for effective translation in the host. Higher stability consequently leads to higher expression rate.

Physico-chemical parameters analysis by ProtParam software indicated that estimated half-life of predicted vaccine was 30 hours in mammalian reticulocytes, more than 20 hours in yeast and more than

10 hours in *E. coli*. In addition, ExPASy ProtParam classifies the chimeric protein as stable protein with instability index, 27.37. Furthermore, the tendency of vaccine solubility upon overexpression in *E. coli* was calculated 0.99% showing an acceptable percentage of solubility in an overexpressed state.

Phyre2 program was used to generate the initial 3D structure of protein vaccine. Obtained results showed that, 33% of residues modelled at > 90% confidence. The evaluation of the initial model with Ramachandran plot showed that the model requires the refinement process. The refinement of modeled structure was carried out in GalaxyRefine server by molecular dynamics simulation. Ramachandran plot assessment after refinement showed refinement applying GalaxyRefine server resulted in modeling of a high-quality 3D model. The analysis of all continuous and discontinuous B-cell epitopes demonstrated that the identified epitopes on protein surface could interact easily with antibodies, and they were generally flexible.

To determine disorder and unstructured regions in the constructed vaccine different server including DisEMBL 1.5, IUPred and DISOPRED3 were used. Obtained results demonstrated that disorder regions were placed between positions 41–71, 361–381. An intrinsically disordered protein (IDP) is classically defined as a protein with lacking a fixed or well-structured three-dimensional fold. IDPs include a spectrum of states from fully unstructured to partially structured and include random coils, (pre-) molten globules, and large multi-domain proteins connected by flexible linkers. Disordered regions are often functional. Many disordered segments fold on binding to their biological targets and some others constitute flexible linkers leading to



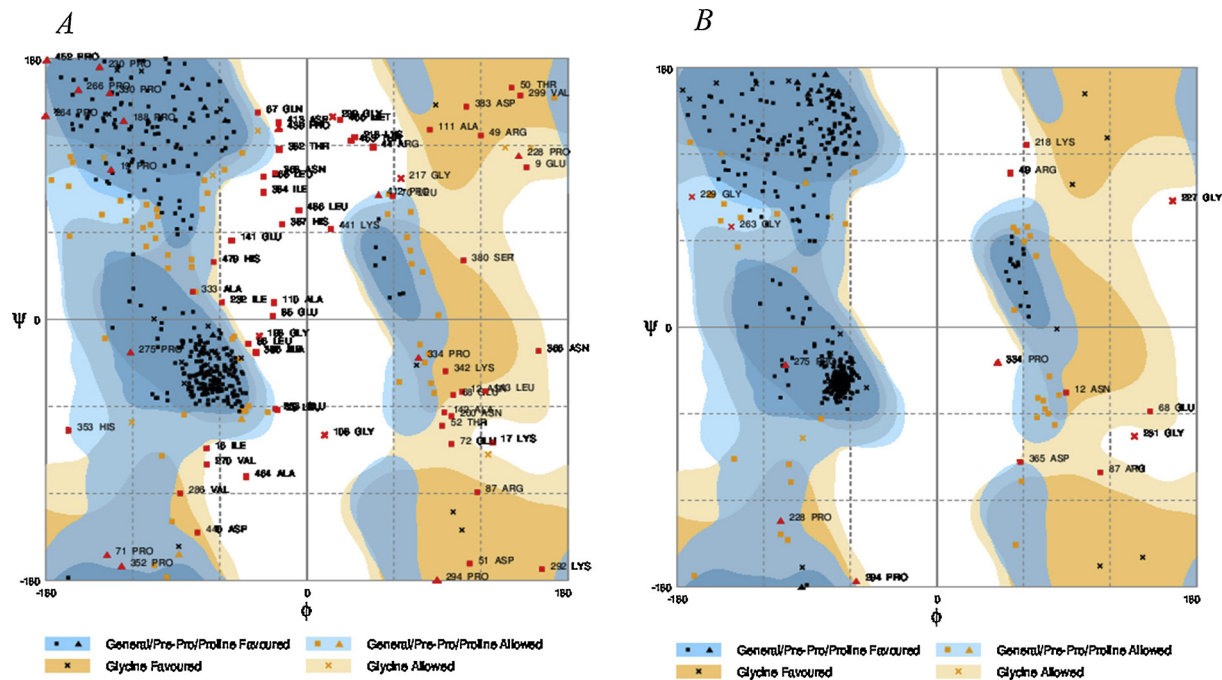


Fig. 9. Validation of protein 3D model, before and after refinement by Ramachandran plot. (a) In initial model, 356 (73.0%), 62 (12.7%) and 70 (14.3%) of residues were located in favored, allowed and outlier regions, respectively. (b) In refined model, in refined model, 435 (89.1%), 39 (8.0%) and 14 (2.9%) of residues were located in favored, allowed and outlier regions, respectively.



Fig. 10. Intrinsically disorder regions. Amino acids in the input sequence were considered disordered when the red line is above the grey dashed line, that is the confidence score is higher than 0.5.

Table 4  
Linear B-Cell epitopes from full length constructed vaccine by using Bcepred.

Position	Epitope	Score
259	QYGTGPGPLRVGYNAPGIP	1
181	KAAAGPGPGNDIISDFGN	1
216	NGKAIHLVNNEGPGPGIEYN	1
288	LRDLKPGPGYVSIDKFRIF	0.999
42	RERAEETRTDTRSRVEESRA	0.996
325	SAAYKPGQAPDGLVDGNKKA	0.98
365	DNSNSDITDQVTWDSNTDI	0.973
464	AQMVKEVSTKTNDVAHEYGA	0.964
400	KLVVVGCFEAERYPDNIHSEA	0.954
439	PDKQKEHAAYFGAPTTTKDG	0.943
63	LTKLQEDLPEQLTELREKFT	0.745

assembly of macromolecular array (Dunker et al., 2001). Taken together, however computational approaches are important step and tools for assessment of the vaccines before initiating the

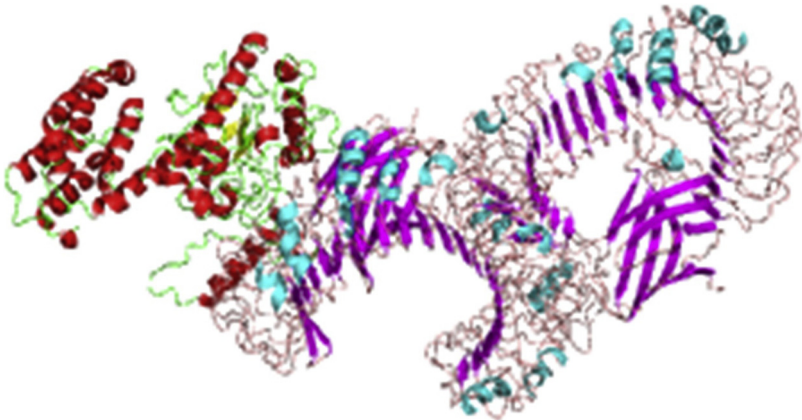
experimental study, constructed vaccine using these immunoinformatics methods should be experimentally evaluated to clarify efficiency and success of final designed vaccine. Therefore, to complete our study on epitope based vaccine development against leptospirosis, it can be suggested in vitro synthesis and in vivo experimental studies to assess the different physical and chemical aspect of this multi-epitope vaccine.

5. Conclusion

In the present study, we design a multi-epitope vaccine candidate based on epitopes of Hap1, LigA, LAg42, SphH and HSP58 antigens employing computational servers. Immunoinformatics analyses indicated that this vaccine is a qualified candidate for immune response reinforcement against leptospirosis by inducing IFN-gamma production. Thus, this multi-epitope vaccine can possibly be utilized for prophylactic or therapeutic usages.

**Table 5**  
Epitopes predicted in constructed by applying Bcepred software based on different parameters.

Prediction parameter	Epitope sequence
Hydrophilicity	KMAENSNIDD, RERAETRTDTRSRVEESR, SQQSFEI, GPGPGND, VNNEGPGRP, GTGPGPG, KPGQAPDG, VDGNNKAY, ERAGDNSNSDITDQ, WDSSTNDI, AERYPDN, PDKQKEHAA, KEVSTKTNDVA
Flexibility	TNLRERAETRTDTRSRVEESR, AALERLRSQQ, SDISGFN, HLVNNEG, RDLKGGP, DGLVDGNK, ALERAGDNSN, VTWSSN, EALERAG, NCLPDKQK, QMVKEVSTKT
Accessibility	KMAENSNIDD, ITNLRERAETRTDTRSRVEESRRLTKLQEDLPEQLTELREKFTAELRKAAGY, EAATSRYNELVER, LERLRSQQSFEIV, NNEGPGP, EYNDMFNN, EQYGTGP, PLYKKMEAVKLRDLKGGP, KALNPKEIEKLYTSY, AAYKPGQAPD, VDGNNKAYLY, PHEYGAE, ERAGDNSNSDITDQ, TWSSSTND, LERAGQKL, FAERYPDNIHSEAAYS, NCLPDKQKEHAAY, KAQMVKVSTKTNDVA
Turns	AENSNIDD, DMFNNFTV, AGDNSNSDIT, TWSSSTNDI
Exposed Surface	RERAETRTDTRSR, ERLRSQQ, PLYKKMEAVK, KALNPKEIEKLY, CLPDKQKEHAA
Polarity	ITNLRERAETRTDTRSRVEESRRLTKLQE, EQLTELREKFTAELRKAAGY, RYNELVERGEAALERLR, QQSFEIVSARAE, TRAVGERA, PLYKKMEAVKLRDLKG, KFRIFCKALNPKEIEKLYT, PHEYGAEA, SHEYGAEALERAGQKL, HSEAAAYK, LPDKQKEHAAYF, KAQMVKVSTKT, AHEYGAEA
Antigenic Propensity	KLVGIELP, SSVITYP, FTVSFWLRVPKVS, KLYTSYLS, YYLYVWIPHEY, GQKLVVVGCF, SLIGFYCLFMFFLNCLPD



**Fig. 11.** Docking model of TLR4/MD-2 and vaccine complex.

**Conflict of interest**

Authors declare no conflict of interests.

**Acknowledgements**

This study was supported by Deputy for Research and Technology of Babol University of Medical Sciences (Grant No. 4869).

**References**

Adler, B., 2015. Vaccines against leptospirosis. *Curr Top Microbiol Immunol* 387, 251–272.

Adler, B., de la, Pena, Moctezuma, A., 2010. Leptospira and leptospirosis. *Vet Microbiol* 140, 287–296.

Arai, R., Ueda, H., Kitayama, A., Kamiya, N., Nagamune, T., 2001. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein engineering* 14, 529–532.

Bhasin, M., Raghava, G.P., 2004. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine* 22, 3195–3204.

Branger, C., Chatrenet, B., Gauvrit, A., Aviat, F., Aubert, A., Bach, J.M., Andre-Fontaine, G., 2005. Protection against Leptospira interrogans sensu lato challenge by DNA immunization with the gene encoding hemolysin-associated protein 1. *Infection and immunity* 73, 4062–4069.

Branger, C., Sonrier, C., Chatrenet, B., Klonjowski, B., Ruvoen-Clouet, N., Aubert, A., Andre-Fontaine, G., Eloit, M., 2001. Identification of the hemolysin-associated protein 1 as a cross-protective immunogen of Leptospira interrogans by adenovirus-mediated vaccination. *Infection and immunity* 69, 6831–6838.

Bui, H.H., Sidney, J., Li, W., Füsseder, N., Sette, A., 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics* 8, 361.

Chaplin, D.D., 2010. Overview of the immune response. *The Journal of allergy and clinical immunology* 125, S3–23.

Christova, I., Tasseva, E., Manev, H., 2003. Human leptospirosis in Bulgaria, 1989–2001: epidemiological, clinical, and serological features. *Scand J Infect Dis* 35, 869–872.

Dhanda, S.K., Vir, P., Raghava, G.P., 2013. Designing of interferon-gamma inducing MHC class-II binders. *Biology direct* 8, 30.

Doytchinova, I.A., Flower, D.R., 2007. VaxiJen: a server for prediction of protective

antigens, tumour antigens and subunit vaccines. *BMC bioinformatics* 8, 4.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., Ausio, J., Nissen, M.S., Reeves, R., Kang, C., Kissinger, C.R., Bailey, R.W., Griswold, M.D., Chiu, W., Garner, E.C., Obradovic, Z., 2001. Intrinsically disordered protein. *Journal of molecular graphics & modelling* 19, 26–59.

Evangelista, K.V., Coburn, J., 2010. Leptospira as an emerging pathogen: a review of its biology, pathogenesis and host immune responses. *Future microbiology* 5, 1413–1425.

Flannery, B., Costa, D., Carvalho, F.P., Guerreiro, H., Matsunaga, J., Da Silva, E.D., Ferreira, A.G., Riley, L.W., Reis, M.G., Haake, D.A., Ko, A.I., 2001. Evaluation of recombinant Leptospira antigen-based enzyme-linked immunosorbent assays for the serodiagnosis of leptospirosis. *Journal of clinical microbiology* 39, 3303–3310.

Fraga, T.R., Barbosa, A.S., Isaac, L., 2011. Leptospirosis: aspects of innate immunity, immunopathogenesis and immune evasion from the complement system. *Scand J Immunol* 73, 408–419.

Fraser, C.K., Diener, K.R., Brown, M.P., Hayball, J.D., 2007. Improving vaccines by incorporating immunological adjuvants. *Expert review of vaccines* 6, 559–578.

Grote, A., Hiller, K., Scheer, M., Munch, R., Nortemann, B., Hempel, D.C., Jahn, D., 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 33, W526–531.

Jones, D.T., Cozzetto, D., 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics (Oxford, England)* 31, 857–863.

Karkhah, A., Amani, J., 2016. A potent multivalent vaccine for modulation of immune system in atherosclerosis: an in silico approach. *Clin Exp Vaccine Res.* 5, 50–59.

Karkhah, A., Saadi, M., Nouri, H.R., 2017. In silico analyses of heat shock protein 60 and calreticulin to designing a novel vaccine shifting immune response toward T helper 2 in atherosclerosis. *Comput Biol Chem* 67, 244–254.

Katz, A.R., Ansdell, V.E., Effler, P.V., Middleton, C.R., Sasaki, D.M., 2001. Assessment of the clinical presentation and treatment of 353 cases of laboratory-confirmed leptospirosis in Hawaii, 1974–1998. *Clin Infect Dis* 33, 1834–1841.

Kelley, L.A., Sternberg, M.J., 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols* 4, 363–371.

Koizumi, N., Watanabe, H., 2003. Identification of a novel antigen of pathogenic Leptospira spp. that reacted with convalescent mice sera. *Journal of medical microbiology* 52, 585–589.

Koizumi, N., Watanabe, H., 2004. Leptospiral immunoglobulin-like proteins elicit protective immunity. *Vaccine* 22, 1545–1552.

Koizumi, N., Watanabe, H., 2005. Leptospirosis vaccines: past, present, and future. *J Postgrad Med* 51, 210–214.

Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS*

- computational biology 8, e1002829.
- Lee, S.H., Kim, S., Park, S.C., Kim, M.J., 2002. Cytotoxic activities of *Leptospira* interorgans hemolysin SphH as a pore-forming protein on mammalian cells. *Infection and immunity* 70, 315–322.
- Lin, Y.P., Lee, D.W., McDonough, S.P., Nicholson, L.K., Sharma, Y., Chang, Y.F., 2009. Repeated domains of leptospira immunoglobulin-like proteins interact with elastin and tropoelastin. *The Journal of biological chemistry* 284, 19380–19391.
- Lin, Y.P., McDonough, S.P., Sharma, Y., Chang, Y.F., 2010. The terminal immunoglobulin-like repeats of LigA and LigB of *Leptospira* enhance their binding to gelatin binding domain of fibronectin and host cells. *PloS one* 5, e11301.
- Livingston, B., Crimi, C., Newman, M., Higashimoto, Y., Appella, E., Sidney, J., Sette, A., 2002. A rational strategy to design multi-epitope immunogens based on multiple Th lymphocyte epitopes. *Journal of immunology* 168, 5499–5506.
- Lovell, S.C., Davis, I.W., Arendall 3rd, W.B., de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2003. Structure validation by Alpha geometry: phi, psi and Cbeta deviation. *Proteins* 50, 437–450.
- Magnan, C.N., Randall, A., Baldi, P., 2009. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics (Oxford, England)* 25, 2200–2207.
- McGuffin, L.J., Bryson, K., Jones, D.T., 2000. The PSIPRED protein structure prediction server. *Bioinformatics (Oxford, England)* 16, 404–405.
- Merien, F., Baranton, G., Perolat, P., 1997. Invasion of Vero cells and induction of apoptosis in macrophages by pathogenic *Leptospira* interorgans are correlated with virulence. *Infect Immun* 65, 729–738.
- Meza, B., Ascencio, F., Sierra-Beltran, A.P., Torres, J., Angulo, C., 2017. A novel design of a multi-antigenic, multistage and multi-epitope vaccine against *Helicobacter pylori*: An in silico approach. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* 49, 309–317.
- Nezafat, N., Karimi, Z., Eslami, M., Mohkam, M., Zandian, S., Ghasemi, Y., 2016. Designing an efficient multi-epitope peptide vaccine against *Vibrio cholerae* via combined immunoinformatics and protein interaction based approaches. *Computational biology and chemistry* 62, 82–95.
- Palaniappan, R.U., McDonough, S.P., Divers, T.J., Chen, C.S., Pan, M.J., Matsumoto, M., Chang, Y.F., 2006. Immunoprotection of recombinant leptospiral immunoglobulin-like protein A against *Leptospira* interorgans serovar Pomona infection. *Infection and immunity* 74, 1745–1750.
- Palaniappan, R.U., Ramanujam, S., Chang, Y.F., 2007. Leptospirosis: pathogenesis, immunity, and diagnosis. *Curr Opin Infect Dis* 20, 284–292.
- Panina-Bordignon, P., Tan, A., Termijtelen, A., Demotz, S., Corradin, G., Lanzavecchia, A., 1989. Universally immunogenic T cell epitopes: promiscuous binding to human MHC class II and promiscuous recognition by T cells. *European journal of immunology* 19, 2237–2242.
- Park, H., Seok, C., 2012. Refinement of unreliable local regions in template-based protein models. *Proteins* 80, 1974–1986.
- Park, S.H., Ahn, B.Y., Kim, M.J., 1999. Expression and immunologic characterization of recombinant heat shock protein 58 of *Leptospira* species: a major target antigen of the humoral immune response. *DNA and cell biology* 18, 903–910.
- Saadi, M., Karkhah, A., Nouri, H.R., 2017. Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infect Genet Evol* 51, 227–234.
- Saha, S., Raghava, G.P., 2006a. AllPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids research* 34, W202–209.
- Saha, S., Raghava, G.P., 2006b. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65, 40–48.
- Schmid, G.P., Steere, A.C., Kornblatt, A.N., Kaufmann, A.F., Moss, C.W., Johnson, R.C., Hovind-Hougen, K., Brenner, D.J., 1986. Newly recognized *Leptospira* species (*"Leptospira inadai"* serovar lyme) isolated from human skin. *J Clin Microbiol* 24, 484–486.
- Schnoor, M., Alcaide, P., Voisin, M.-B., van Buul, J.D., 2015. Crossing the vascular wall: common and unique mechanisms exploited by different leukocyte subsets during extravasation. *Mediators of inflammation* 2015.
- Schoenborn, J.R., Wilson, C.B., 2007. Regulation of interferon-gamma during innate and adaptive immune responses. *Adv Immunol* 96, 41–101.
- Sen, T.Z., Jernigan, R.L., Garnier, J., Kloczkowski, A., 2005. GOR V server for protein secondary structure prediction. *Bioinformatics (Oxford, England)* 21, 2787–2788.
- Silva, E.F., Medeiros, M.A., McBride, A.J., Matsunaga, J., Esteves, G.S., Ramos, J.G., Santos, C.S., Croda, J., Homma, A., Dellagostin, O.A., Haake, D.A., Reis, M.G., Ko, A.I., 2007. The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. *Vaccine* 25, 6277–6286.
- Tourani, M., Karkhah, A., Najafi, A., 2017. Development of an epitope-based vaccine inhibiting immune cells rolling and migration against atherosclerosis using in silico approaches. *Comput Biol Chem* 70, 156–163.
- Trueba, G., Zapata, S., Madrid, K., Cullen, P., Haake, D., 2004. Cell aggregation: a mechanism of pathogenic *Leptospira* to survive in fresh water. *Int Microbiol* 7, 35–40.
- Vinetz, J.M., 2001. Leptospirosis. *Curr Opin Infect Dis* 14, 527–538.
- Waitkins, S.A., 1986. Leptospirosis as an occupational disease. *Br J Ind Med* 43, 721–725.
- Wang, Z., Jin, L., Wegrzyn, A., 2007. Leptospirosis vaccines. *Microb Cell Fact* 6, 39.
- Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research* 35, W407–410.
- Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology (Clifton, N.J.)* 112, 531–552.
- Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. *BMC bioinformatics* 9, 40.
- Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* 31, 3406–3415.